In the specification:

Please substitute the following paragraphs:

Page 5, lines 6-25:

(60/263,668) Another major cost of insulin production is purification. Chromatography accounts for 30% of operating expenses and 70% of equipment in production of insulin (Petridis ~~et al.~~ *et al.*, 1995). Therefore, new approaches are needed to minimize ir eliminate chroma-tography in insulin production. One such approach is the use of GVGVP (SEQ ID NO. 1) as a fusion protein to facilitate single step purification without the use of chromatography. GVGVP (SEQ ID NO. 1) is a Protein Based Polymer (PBP) made from synthetic genes. At lower temperatures this polymer exists as more extended molecules. Upon raising the temperature above the transition range, polymer hydrophobically folds into dynamic structures called β-spirals that further aggregate by hydrophobic association to form twisted filaments (Urry, 1991; Urry ~~et al.~~*et al.*, 1994). Inverse temperature transition offers several advantages. It facilitates scale up of purfication from grams to kilograms. Milder purification condition requires only a modest change in temperature and ionic strength. This should also facilitate higher recovery, faster purification, and high volume processing. Protein purification is generally the slow step (bottleneck) in pharmaceutical product development. Through exploitation of this reversible inverse temperature transition property, simple and inexpensive extraction and purification may be performed. The temperature at which the aggregation takes place can be manipulated by engineering biopolymers containing varying numbers of repeats and changing salt concentration in solution (McPherson ~~et al.~~*et al.*, 1996). Chloroplast mediated expression of insulin-polymer fusion protein should eliminate the need for the expensive fermentation process as well as reagents needed for recombinant protein purification and downstream processing.

Page 10, lines 10-25:

(~~60/,263,668~~60/263,668) A remarkable feature of chloroplast genetic engineering is the observation of exceptionally large accumulation of foreign proteins in transgenic plants. This can be as much as 46% of CRY protein in total soluble protein, even in bleached old leaves (DeCosa ~~et al.~~*et al.*, 2001). Stable expression of a pharmacutical protein in chloroplasts was first reported for

GVGVP (SEQ ID NO. 1), a protein-based polymer with varied medical applications (such as the prevention of post-surgical adhesions and scars, wound coverings, artificial pericardia, tissue reconstruction, and programmed drug delivery) (Guda ~~et al.~~*et al.*, 2000). Susequently, expression of the human somatotropom via the tobacco chloroplast genome (Staub ~~et al.~~*et al.*, 2000) to high levels (7% of total soluble protein) was observed. The following investigations that are in process illustrate the power of this technology to express small peptides, entire operons, vaccines that require oligomeric proteins with stable disulfide bridges, and monoclonals that require assembly of heavy/light chains via chaperonins. It is essential to develop a selection system free of antibiotic resistant genes for the edible insulin approach to be successful. One such marker free chloroplast transformation system has been accomplished (Daniell ~~et al.~~*et al.*, 2000). Experiments are in progress to develop chloroplast transformation of edible leaves (alfalfa and lettuce) for the practical applications of this approach.

Page 11, lines 7-22:

(60,185,987) GVGVP (SEQ ID NO. 1) is a PBP made from synthetic genes. At lower temperatures the polymers exist as more extended molecules which, on raising the temperature above the transition range, hydrophobically fold into dynamic structures called β-spirals that further aggregate by hydrophobic association to form twisted filaments (Urry, 1991; Urry ~~et al.~~*et al.*, 1994). Inverse temperature transition offers several advantages. Expense associated with chromatographic resins and equipment are eliminated. It also facilitates scale up of purification from grams to kilograms. Milder purification conditions use only a modest change in temperature and ionic strength. This also facilitates higher recovery, faster purification, and high volume processing. Protein purification is generally the slow step (bottleneck) in pharmaceutical product development. Through exploitation of this reversible inverse temperature transition property, simple and inexpensive extraction and purification may be performed. The temperature at which the aggregation takes place can be manipulated by engineering biopolymers containing varying numbers of repeats and changing salt concentration in solution (McPherson ~~et al.~~*et al.*, 1996). Chloroplast mediated expression of insulin-polymer fusion protein should eliminate the need for the expensive fermentation process as well as reagents needed for recombinant protein purification and downstream processing.

Page 12, lines 12-31:

(60/185,987) In accordance with one advantageous feature of this invention, we use poly(GVGVP) as a fusion protein to enable hyper-expression of insulin and accomplish rapid one -step purification of fusion peptides utilizing the inverse temperature transition properties of this polymer. In another advantageous feature of this invention, we develop insulin-CTB fusion protein for oral delivery in nicotine-free edible tobacco (LAMD 605). Both features are accomplished as follows:

a) Develop recombinant DNA vectors for enhanced expression of Proinsulin as fusion proteins with GVGVP (SEQ ID NO. 1) or CTB via chloroplast genomes of tobacco,

b) Obtain transgenic tobacco (Petit Havana and LAMD 605) plants,

c) Characterize transgenic expression of proinsulin polymer or CTB fusion proteins using molecular and biochemical methods in chloroplasts,

d) Employ existing or modified methods of polymer purification from transgenic leaves,

e) Analyze Mendelian or maternal inheritance of transgenic plants,

f) Large scale purification of insulin and comparison of current insulin purification methods with polymer-based purfication method in *E. coli* and tobacco,

g) Compare natural refolding chloroplasts with *in vitro* processing,

h) Characterization (yield and purity) of proinsulin produced *E. coli* and transgenic tobacco, and

i) Assessment of diabetic symptoms in mice fed with edible tobacco expressing CTB-insulin fusion protein.

Page 14, lines 14-29:

(60/185,987) Protein Based Polymers (PBP): The synthetic gene that codes for a bioelastic PBP was designed after repeated amino acid sequences GVGVP (SEQ ID NO. 1), observed in all sequenced mammalian elastin proteins (Yeh ~~et al.~~*et al.*, 1987). Elastin is one of the strongest known natural fibers and is present in skin, ligaments, and arterial walls. Bioelastic PBPs containing multiple repeats of this pentamer have remarkable elastic properties, enabling several medical and non-medical applications (~~Urry et al. 1993, Urry 1995, Daniell 1995~~Urry *et al.*, 1993; Urry, 1995; Daniell, 1995). GVGVP (SEQ ID NO. 1) polymers prevent adhesions following surgery, aid in reconstructing tissues, and delivering drugs to the body over an extended period of time. North

American Science Associates, Inc. reported that GVGVP (SEQ ID NO. 1) polymer is non-toxic in mice, non-sensitizing and non-antigenic in guinea pigs, and non-pyrogenic in rabbits (Urry ~~et al.~~*et al.*, 1993). Researchers have also observed that inserting sheets of GVGVP (SEQ ID NO. 1) at the sites of contaminated wounds in rats reduces the number of adhesions that form as the wounds heal (Urry ~~et al.~~*et al.*, 1993). In a similar manner, using the GVGVP (SEQ ID NO. 1) to encase muscles that are cut during eye surgery in rabbits prevents scarring following the operation (~~Urry et al. 1993, Urry 1995~~ ~~1995~~Urry *et al.*, 1993; Urry, 1995). Other medical applications of bioelastic PBPs include tissue reconstruction (synthetic ligaments and arteries, bones), wound coverings, artificial pericardia, catheters, and programmed drug delivery (Urry, 1995; Urry, ~~et al.~~*et al.*, 1993, 1996).

Pages 20-21, lines 20-39:

(60/185,987) Recently, the human pre-proinsulin gene was obtained from Genentech, Inc. First, the pre-proinsulin was sub-cloned into pUC19 to facilitate further manipulations. The next step was to design primers to make chloroplast expression vectors. Since we are interested in proinsulin expression, the 5′ primer was designed to land on the proinsulin sequence. This FW primer eluded the 69 bases or 23 coded amino acids of the leader or pre-sequence of preproinsulin. Also, the forward primer included the enzymatic cleavage site for the protease factor Xa to avoid the use of cyanogen bromide. Beside the Xa-factor, a SmaI site was introduced to facilitate subsequent subcloning. The order of the FW primer sequence is smal-Xa-factor-Proinsulin gene. The reverse primer includes BamHI and XbaI sites, plus a short sequence with homology with the pUC19 sequence following the proinsulin gene. The 297bp PCR produce (Xa Pris) includes three restriction sites, which are the SmaI site at the 5′-end and XbaI/BamHI sites at the 3′end of the proinsulin gene. The Xa-Pris was cloned into pCR2.1 resulting in pCR2.1-Xa-Pris (4.2kb). Insertion of Xa-Pris into the multiple cloning site of pCR2.1, resulted in additional flanking restriction enzyme sites that will be used in subsequent sub-cloning steps. A GVGVP (SEQ ID NO. 1) 50-mer was generated as described previously (Daniell ~~et al.~~*et al.*, 1997). The ribosome binding sequence was introduced by digesting pUCs-10, which contains the RBS sequence GAAGGAG, with NooI and Hind III flanking sites. The plasmid pUC19-50 was also digested with the same enzymes. The 50mer gene was eluted from the gel and ligated to pUCs-10 to produce pUCs-10-50mer. The ligation step inserted into the 50mer gene a RBS sequence and a SmaI site outside the gene to facilitate subsequent fusion to proinsulin.

Page 25, lines 15-30:

(60,185,987) Protease Xa Digestion of the Biopolymer-proinsulin fusion protein and Purification of Proinsulin: Factor Xa was purchased from New England Biolabs at a concentration of 1.0 mg/ml. The Factor Xa is supplied in 20mM HEPES, 500mM, NaCl, 2mM $CaCl_2$, 50% glycerol, (pH 8.0). The reaction was carried out in a 1:1 ratio of fusion protein to reaction buffer. The reaction buffer was made with 20mM Tris-HCl, 100mM NaCl, 2mM $CaCl_2$, (pH 8.0). The enzymatic cleavage of the fusion protein to release the proinsulin protein from the $(GVVP)_{40}$ was initiated by adding the protease to the purified fusion protein at a ratio (ww) of approximately 1,500. This digestion was continued for 5 days with mild stirring at 4°C. Cleavage of the fusion protein was monitored by SDS-PAGE analysis. After the cleavage, the same conditions are used for purification of the fusion protein, except that instead of recovering the pellet, the supernatant is saved. We detected cleaved proinsulin in the extracts isolated in 6M guanidine hydrochloride buffer as shown in Fig. 1C 11. Conditions can be estimized for complete cleavage. The Xa protease has been successfully used to cleave $(GVGVP)_{20}$-GST fusion (McPherson ~~et al.~~*et al.*, 1992). Therefore, cleavage of proinsulin from GVGVP (SEQ ID NO. 1) using the Xa protease does not pose problems.

Page 29, lines 6-23:

(60/185,987) CTB-Proinsulin Vector Construction: The chloroplast expression vector pLD-CTB-Proins was constructed as follows. First, both proinsulin and cholera toxin B-subunit genes were amplified from suitable DNA using primer sequences. Primer 1 contains the GGAGG chloroplast preferred ribosome binding site five nucleotides upstream of the start codon (ATG) for the CTB gene and a suitable restriction enzyme site (SpeI) for insertion into the chloroplast vector. Primer 2 eliminates the stop codon and adds the first two amino acids of a flexible hinge tetrapeptide GPGP (SEQ ID NO. 3) as reported by Bergerot ~~et al.~~*et al.*, (1997), in order to facilitate folding of the CTB-proinsulin fusion protein. Primer 3 adds the remaining two amino acids for the hinge tetrapeptide and eliminates the pre-sequence of the pre-proinsulin. Primer 4 adds a suitable restriction site (SpeI) for subcloning into the chloroplast vector. Both the CTB and proinsulin PCR fragments were excised at the SmaI and XbaI restriction sites. Eluted fragments were ligated into the TA cloning vector. Interestingly, all white colonies showed the wrong orientation of the CTB insert. The CTB-proinsulin fragment was excised at the EcoRI sites and inserted into EcoRI digested dephosphorolated pLD vector. Resultant onicroplast integration expression vector, pLD-CTB-Proins

will be tested for expression in *E.coli* by western blots. After confirmation of expression of CTB-proinsulin fusion in *E.coli*, pLD-CTB-Proins will be bombarded into tobacco cells as described below.

Page 29, line 24 through Page 30, line 14:

(60/263,668) The following vectors may be designed to optimize protein expression, purification, and production of proteins with the same amino acid composition as in human insulin.

a) Using tobacco plants, Eibl (1999) demonstrated, *in vivo*, the differences in translation efficiency and mRNA stability of a GUS reporter gene due to various 5′ and 3′ untranslated regions (UTR's). This already described systematic transcription and translation analysis can be used in a practical endeavor of insulin production. Consistent with Eibl's (1999) data for increased translation efficiency and mRNA stability, the psbA 5′ UTR can be used in addition with the psbA 3′ UTR already in use. The 200 bp tobacco chloroplast DNA fragment containing 5′ psbA UTTR may be amplified by PCR using tobacco chloroplast DNA as template. This fragment may be cloned directly in the pLD vector multiple cloning site downstream of the promoter and the aadA gene. The cloned sequence may be exactly the same as in the psbA gene. (Update "Human Insulin") We have cloned the 5′ untranslated region of the tobacco psbA gene including the promoter (5′UTR), shown in Figure 32. We performed PCR using the primers CCGTCGACGTAGAGAAGTCCGTATT (SEQ ID NO. 4) and GCCCATGGTAAAATCTTGGTTTATTTA (SEQ ID NO. 5), which resulted in a 200 base pair product, as expected. We inserted this PCR product into a TA cloning vector. Since restriction enzyme sites were not available to subclone the 5′UTR immediately upstream of the gene coding for the CTB-proinsulin fusion protein, we used the "SOEing" PCR technique to create the DNA sequence with the 5′UTR immediately upstream of the CTB-proinsulin gene (Figure 33). The products of this PCR include both the 5′UTR (200bp) and the gene for CTB-proinsulin (600bp) as additional products as well as the desired 5′UTR CTB-proinsulin (5CP) at 800 bp. 5CP was eluted and then inserted into the TA cloning vector where DNA sequencing was performed to confirm accuracy of nucleotide sequence before it was subcloned into the pLD vector.

Page 34, lines 1-18:

(60/185,987) Optimization of fusion gene expression: It has been reported that foreign genes are expressed between 5% (cry1AC, cry11A) and 30% (uldA) in transgenic chloroplasts (Daniell, 1999). If the expression levels of the CTB-Proinsulin or polymer-proinsulin fusion proteins are low, several approaches will be used to enhance translation of these proteins. In chloroplast, transcription regulation of gene expression is less important, although some modulations by light and developmental conditions are observed (Cohen and Mayfield, 1977). RNA and protein stability appear to be less important because of observation of large accumulation of foreign proteins (e.g. GUS up to 30% of total protein) and tps1 transcripts 16,966-fold higher than the highly expressing nuclear transgenic plants. Chloroplast gene expression is regulated to a large extent at the post-transcriptional level. For example, 5'UTRs are used for optional translation of chloroplast mRNAs. Shine-Delgarno (GGAGG) sequences as well as a stem-loop structure located 5'adjacent to the SD sequence are used for efficient translation. A recent study has shown that insertion of the psbA 5'UTR downstream of the 16S rRNA promoter enhanced translation of a foreign gene (GUS) hundred-fold (Eibl et al. 1999). Therefore, the 85-bp tobacco chloroplast DNA fragment (1595-1680) containing 5'psbA UTR will be amplified using the following primers cctttaaaaagccttccatttctattt gccatggtaaaatcttggtttatta (SEQ ID NO. 6 and SEQ ID NO. 7, respectively). This PCR product will be inserted downstream of the 16S rRNA promoter to enhance translation of the proinsulin fusion proteins.

Page 34, line 19 through page 35, line 4:

(60/185,987) Yet another approach for enhancement of translation is to optimize codon compositions of these fusion protein. Since both fusion proteins are expressed well in *E. coli*, we expected efficient expression in chloroplasts. However, optimizing codon compositions of proinsulin and CTB genes to ~~march~~match the psbA gene could further enhance the level of translation. Although rbcL (RuBisCO) is the most abundant protein on

earth, it is not translated as frequently as the psbA gene due to the extremely high turnover of the psbA gene product. The psbA gene is under stronger selection for increased translation efficiency and is the most abundant thylakoid protein. In addition, codon usage in higher plant chloroplasts is biased towards the NNC codon of 2-fold degenerate groups (i.e. TTC over TTT, GAC over GAT, CAC over CAT, AAC over AAT, ATC over ATT, ATA etc.) This is in addition to a strong bias towards T at third position of 4-fold degenerate groups. There is also a context effect that should be taken into consideration while modifying specific codons. The 2-fold degenerate sites immediately upstream from a GNN codon do not show this bias towards NNC, (TTT GGA is preferred to TTC GGA while TTC CGT is preferred to TTT CGT TTC AGT to TTT AGT and TTC TCT to TTT TCT) SEQ ID NO. 8. In addition, highly expressed chloroplast genes use GNN more frequently than other genes. ~~The web site _____ may be used to optimize codon composition by comparing different species.~~ Abundance of amino acids in chloroplasts can be taken into consideration (pathways compartmentalized in plastids as opposed to those that are important into plastids.)

Page 40, line 21 through Page 41, line 5:

(60/185,987) After electrophoresis, proteins are transferred to a nitrocellulose membrane electrophoretically in 25 mM Tris, 192mM glycine, 5% methanol (pH 8.3). The filter is blocked with 2% dry milk in Tris-buffered saline for two hours at room temperature and stained with antiserum raised against the polymer AVGVP (SEQ ID NO. 9) (kindly provided by the University of Alabama at Birmingham, monoclonal facility) overnight in 2% dry milk/Tris buffered saline. The protein bands reacting to the antibodies are visualized using alkaline phosphatase-linked secondary antibody and the substrates nitroblue tetrazolium and 5-bromo-4-chloro-3-indolyl-phosphate (Bio-Rad). Alternatively for insulin-polymer fusion proteins, a Mouse anti-human proinsulin (IgGl) monoclonal antibody is used as a primary antibody. To detect the binding of the primary antibody to the recombinant proinsulin, a Goat anti-mouse IgG Horseradish Peroxidase Labeled monoclonal antibody (HPR) is used. The substrate used for conjugation with HPR is 3,3′, 5,5′-Tetramethylbenzidine. All products are available from American Qualex Antibodies, San Clemente, CA. As a positive control, human recombinant proinsulin from Sigma may be

used. This human recombinant proinsulin was expressed in *E. coli* by a synthetic proinsulin gene. Quantification of purified polymer fusion proteins is carried out by densitometry using Scanning Analysis software (BioSoft, Ferguson, MO) installed in a Macintosh LC III computer (Apple Computer, Cupertino, USA) with a 160-Mb hard disk operating on a System 7.1, connected by SCSI interface to a Relisys RELI 2412 Scanner (Relisys, Milpitas, CA). Total protein contents is then determined by the dye-binding assay using reagents supplied in kit ~~fro~~ from Bio-Rad, with bovine serum albumin as a standard.

Page 52, lines 17-27:

A remarkable feature of chloroplast genetic engineering is the observation of exceptionally large accumulation of foreign proteins in transgenic plants, as much as 46% of CRY protein in total soluble protein, even in bleached old leaves (3). Stable expression of a pharmaceutical protein in chloroplasts was first reported for GVGVP (SEQ ID NO. 1), a protein based polymer with varied medical applications (such as the prevention of post-surgical adhesions and scars, wound coverings, artificial pericardia, tissue reconstruction and programmed drug delivery (88)). Subsequently, expression of the human somatotropin via the tobacco chloroplast genome (9) to high levels (7% of total soluble protein) was observed. The following investigations that are in progress in the Daniell laboratory illustrate the power of this technology to express small peptides, entire operons, vaccines that require oligomeric proteins with stable disulfide bridges and monoclonals that require assembly of heavy/light chains via chaperonins.

Page 63, lines 6-15:

For IGF-I expression, the use of the TEV protease (Gibco cat n 10127-017) would be ideal. The cleavage site that is recognized for this protease is Glu-Asn-Leu-Tyr-Phe-Gln-Gly (SEQ ID NO. 10) and it cuts between Gln-Gly. This strategy allows the release of the mature protein by incubation with TEV protease leaving a glycine as the first amino acid consistent with human mature IGF-I protein.

The purification system of the *E. coli* Interferon-α5 expression method was based on 6 Histidine-tags that bind to a nickel column and biotinylated thrombin to eliminate the tag (SEQ ID NO. 12) on IFN-α5. Thrombin recognizes Leu-Val-Pro-Arg-Gly-Ser (SEQ ID NO.

11) and cuts between Arg and Gly. This leaves two extra amino acids in the mature protein, but antiviral activity studies have shown that this protein is at least as active as commercial IFN-α2.

Page 78, lines 14-21:

PCR Analysis: Total plant DNA from putative transgenic and untransformed plants was isolated using the Dneasy kit (Qiagen). PCR primers 3P (5'AAAACCCGTCCTCAGTTCGGATTGC-3') (SEQ ID NO. 13) and 3M (5'-CCGCGTTGTTTCATCAAGCCTTAG-3') (SEQ ID NO. 14) were used for PCR on putative transgenic and untransformed plant total DNA. Samples were carried through 30 cycles using the following temperature sequence: 94°C for 1 min, 62°C for 1.5 min and 72°C for 2 min. Cycles were preceded by denaturation for 5 min at 94°C. PCR confirmed shoots from the second selection were transferred to rooting medium (MSO medium containing 500 mg/L spectinomycin).